

# UCSF

## UC San Francisco Previously Published Works

### Title

Sensory transformations and the use of multiple reference frames for reach planning.

### Permalink

<https://escholarship.org/uc/item/294676xx>

### Journal

Nature neuroscience, 12(8)

### ISSN

1097-6256

### Authors

McGuire, Leah MM

Sabes, Philip N

### Publication Date

2009-08-01

### DOI

10.1038/nn.2357

Peer reviewed



Published in final edited form as:

Nat Neurosci. 2009 August ; 12(8): 1056–1061. doi:10.1038/nn.2357.

## Sensory transformations and the use of multiple reference frames for reach planning

Leah M.M. McGuire and Philip N. Sabes

W. M. Keck Center for Integrative Neuroscience, Department of Physiology, and the Neuroscience Graduate Program, University of California, San Francisco, California

### Abstract

The sensory signals that drive movement planning arrive in a variety of “reference frames”, so integrating or comparing them requires sensory transformations. We propose a model where the statistical properties of sensory signals and their transformations determine how these signals are used. This model captures the patterns of gaze-dependent errors found in our human psychophysics experiment when the sensory signals available for reach planning are varied. These results challenge two widely held ideas: error patterns directly reflect the reference frame of the underlying neural representation, and it is preferable to use a single common reference frame for movement planning. We show that gaze-dependent error patterns, often cited as evidence for retinotopic reach planning, can be explained by a transformation bias and are not exclusively linked to retinotopic representations. Further, the presence of multiple reference frames allows for optimal use of available sensory information and explains task-dependent reweighting of sensory signals.

---

Humans use various sensory signals when interacting with the environment. We can reach to pick up a coin we see in front of us or transfer the coin from one hand to another without looking. Using multiple sensory modalities for planning similar movements is potentially problematic, since different sensory signals arrive in different “reference frames”. Specifically, early visual pathways represent stimulus location relative to current gaze location – a *retinotopic* representation, while proprioceptive signals represent hand location relative to the shoulder or trunk – a *body-centered* representation. In order to utilize such signals, some must be *transformed* between reference frames. While sensory transformations may appear mathematically simple, here we show that transformations can incur a cost by adding bias and variability<sup>1, 2</sup> into the transformed signal. Thus, sensory transformations likely influence the flow of information in motor planning circuits.

It has been argued that transforming sensory signals into a common representation would simplify reach planning<sup>3-7</sup>, and many researchers have attempted to characterize this representation. Many psychophysical studies focus on the patterns of reach error, under the assumption that the reference frame for movement planning directly determines the spatial pattern of errors. Such studies argue for both retinotopic<sup>5, 8-11</sup> and hand- or body-centered<sup>12, 13</sup> planning. Studies of primate physiology and human fMRI also find evidence for a range of neural representations for movement planning<sup>14-23</sup>. These disparate results suggest that a single common representation for reach planning is unlikely<sup>22-26</sup>. We argue

that the presence of noisy sensory transformations makes it advantageous to represent movement plans simultaneously in multiple reference frames.

Our experiment examined how the representations of movement plan depend on the available sensory inputs, focusing on the effects of gaze location. A well studied gaze-dependent error is the *retinal eccentricity effect*, where subjects overestimate the distance between the gaze location and a visually peripheral target when pointing to the target<sup>27</sup>. Since these errors are most parsimoniously described as an overshoot in a retinotopic reference frame, they are cited as evidence for retinotopic reach planning<sup>8-10</sup>. We found that the magnitude and direction of gaze-dependent errors depends heavily on the available sensory information. These results were interpreted using a model of movement planning where sensory signals are combined in a statistically principled manner in two separate reference frames. The model provides a novel explanation for these gaze-dependent reach errors: they arise when sensory information about target location is transformed between representations using an internal estimate of gaze direction that is biased toward the target. We thus demonstrate that spatial patterns of reach errors need not directly reflect the reference frame of the underlying neural representation.

## Results

### Measuring gaze-dependent reach errors

We first examined how the pattern of reach errors depends on the sensory signals available during the planning and execution of a movement. Specifically, we manipulated information about target location and initial hand position, the two variables needed to compute a movement vector. Target information was varied by having subjects reach either to visual targets (VIS), proprioceptive targets (the left index finger, PROP), or targets consisting of simultaneous visual and proprioceptive signals (left index finger with visual feedback, VIS +PROP). Information about initial hand position was varied by having subjects reach either with (FB), or without (NoFB) visual feedback of the right (reaching) hand before movement onset, although feedback was never available during the movement. For each of the six resulting trial types, we measured movement errors as subjects reached to nine different target locations with gaze held on one of two fixation points (Supplemental Fig. S1, online).

A comparison of reach endpoints for an example subject at the midline target illustrates how reach errors depend on the available sensory signals (Fig. 1). The mean error changed markedly as a function of gaze location, and these gaze-dependent effects differed across trial types. During VIS and VIS+PROP trials (Fig. 1a–d), reach endpoints were biased *away* from the gaze location (the retinal eccentricity effect), and the magnitude of the effect decreased with increasing sensory information (Fig. 1a vs. 1b–d). In PROP trials, a small bias *toward* gaze location was observed instead (Fig. 1e,f). These patterns were consistent across targets (Supplemental Section 1.1 and Fig. S2, online). In addition to these gaze-dependent effects, there was a gaze-independent bias in reaching that could differ across targets and trial types. While there was a trend toward overshooting the target, the pattern of this bias, measured across targets and trial types in a separate gaze-free trial condition, was idiosyncratic from subject to subject (Supplemental Section 1.2 and Fig. S3, online), making

these patterns difficult to interpret. We therefore focused on the consistent gaze-dependent effects.

In order to isolate the gaze-dependent effects, we analyzed reach errors in polar coordinates about the midpoint of the eyes and subtracted the gaze-free errors (*Methods*, Supplemental Section 1.2 and Fig. S4, online). In examining radial (depth) reach errors we found a tendency to overshoot the target; however these errors did not differ between the two gaze locations (Supplemental Section 1.3 and Fig. S5, online). In examining angular reach errors (Fig. 2) we observed both a general leftward bias (see Supplemental Section 1.2, online) and significant differences as a function of gaze location. Gaze-dependent effects were thus confined to the angular reach errors.

We next examined how these gaze-dependent effects varied with trial type (Fig. 2a–f). The retinal eccentricity effect was observed in VIS/NoFB trials (Fig. 2a): subjects made rightward (positive) reach errors when fixating to the left of the target and leftward (negative) reach errors when fixating to the right of the target. A similar pattern was observed in VIS/FB trials, but with smaller magnitude (Fig. 2b, and see reference8). In contrast, during PROP trials the reach endpoint was *closer* to the fixation point, so that leftward errors were made when fixating to the left of the target (Fig. 2e,f). The gaze-dependent error patterns in the VIS+PROP trial types appear to be a combination of those observed in VIS and PROP trials (Fig. 2c,d). For all trial types, the errors for the two gaze locations align qualitatively when angular error is plotted in a retinotopic reference frame, i.e. as a function of target relative to gaze (Fig. 2, insets). This alignment might appear to support a retinotopic representation for reach planning8-10. However, since these patterns differ markedly across trial types they cannot be readily explained in terms of a fixed retinotopic bias, suggesting the need for a different explanation of these error patterns.

### A Planning Model: Integration Across Reference Frames

We developed a model of reach planning that accounts for the pattern of gaze-dependent errors observed in our data. The model has two key features: the presence of multiple representations for movement planning and a bias in the transformation between those representations.

The model begins with sensory inputs signaling target location, initial hand position, and gaze location. As sensory signals are inherently variable, we modeled them as Gaussian likelihoods of true location given the sensory input, with likelihood variance reflecting the reliability of the sensory modality28. Visual signals arrive in a retinotopic representation and proprioceptive signals arrive in a body-centered representation. Each available signal is then transformed into the “non-native” reference frame (Fig 3a). Subjects' head positions were fixed during the experiment, so this complex nonlinear transformation29 can be approximated simply by adding or subtracting the gaze location, i.e. by convolving their distributions (Fig. 3a, *Methods* and1). However, since the internal estimate of gaze location is also uncertain, this transformation adds variability to the signals. When both sensory modalities are available, the “native” and transformed signals are integrated in both reference frame representations (VIS+PROP target condition, Fig. 3a). Movement vectors are then computed within each representation, by convolving the target and initial hand

distributions (subtraction). Note that because the sensory transformation adds variability, each spatial variable is more reliably represented in one or the other of these representations, depending on the availability and reliability of the relevant visual and proprioceptive signals (Supplementary Section 2.1, online).

The sensory transformations in our model also introduce a bias into the estimates of transformed variables. In particular, we posit that the internal estimate of gaze direction used to transform target location is biased toward the target. We discuss possible origins for this bias below. The bias was modeled as a Bayesian prior on gaze location centered on the target. Since the transformation consists of adding or subtracting gaze location, depending on the direction of the transformation, the prior effectively biases the transformed target estimate either away from or towards the direction of gaze (Fig. 3a). “Native” (untransformed) target representations remain unbiased. Since the gaze prior is centered on the estimated target location, the variance of the prior (Fig. 3b,c), and hence the magnitude of the bias (Fig. 3d), was assumed to scale with the uncertainty of the internal target estimate. Thus, the gaze-dependent errors in the model depend on the relative weighting of the various sensory signals, depending on both the availability and reliability of sensory inputs and on the method of reading out the final movement vector.

We considered three possible output schemes for reading the planned movement vector from the model (Supplemental Fig. S6, online). Either the retinotopic (RET) or body-centered (BODY) representations can each be read out directly. Alternatively, the two movement vector estimates can themselves be combined to form an integrated readout (INTEG). The contributions of the two representations to the INTEG readout depend on their relative reliability. Note that these calculations use the simplifying assumption that the signals being combined are independent (see Supplemental Section 2.2, online for discussion).

Each of the three potential readouts provided a quantitative prediction of reach errors as a function of hand, target, and gaze locations. The only free parameters in the model were the variances of individual sensory inputs and the gaze priors (*Methods*). The values of the proprioceptive variances were based on previously reported values<sup>30</sup>. Four parameters remained: visual variance, gaze variance, and two scaling factors relating the variance of the gaze prior to the variance in visual and proprioceptive target signals. We fit these parameters to the gaze-dependent errors shown in Figure 2 after mean correction (see *Methods*), using least-squares regression (see Supplemental Section 2.3, Table-S1, and Fig. S7 online, for fit values).

### Model Fits of Constant and Variable Reach Errors

We first consider how well the three output models fit the observed patterns of gaze-dependent errors (Fig. 4a–f). When the model was fit with a single-representation readout, RET or BODY, it failed to predict errors for all trial types. This is because only transformed target signals contain a gaze-dependent error, so when the readout was in the “native” reference frame of a unimodal target, no error was predicted (Fig. 4a,b,e,f). While both readouts had errors for bimodal targets (VIS+PROP, Fig. 4c,d), errors in the RET readout are due to the transformed proprioceptive signal, and would thus be in the wrong direction (Fig 3a). Thus, for the RET readout the fitting procedure nullified the effect of the

transformed proprioceptive signal by forcing the model to rely only on vision, i.e. by driving the visual variance parameter toward zero (Supplemental Table S1, online). In contrast, when the movement vector representations were combined in the INTEG readout, the model performed well across all trial types. This readout captured both changes in the magnitude of gaze-dependent errors (VIS vs. VIS+PROP, and FB vs. NoFB) and the sign reversal observed with PROP targets. It accomplished this by differentially weighting the two reference frames across trial conditions, with both representations making substantive contributions (Supplemental Section 2.4 and Fig. S8, online). Indeed, a scheme that simply switches between the RET and BODY readouts depending on task would not capture the data well. First, both the RET and BODY readouts failed to capture the differences in the magnitude of gaze-dependent errors that we observed between FB and NoFB conditions. Second, in order to predict the observed errors, the switching scheme would need to rely predominantly on the more variable “transformed” signals, rather than “native” signals, a sub-optimal arrangement.

In addition to fitting the gaze dependent error patterns, the model predicted the differences in movement variability across trial types (Fig. 5a). Since computations within the model were assumed to be noise-free, model output variability was due entirely to variability in the sensory inputs, shaped by the model computations, and did not require any additional parameter fitting. The INTEG model fit provided an accurate prediction for the changes in output variance across trial types, better than the two single-representation fits. These predictions came from separate parameter fits for each readout. However the model parameters presumably reflect actual variances in the neuronal representations of sensory inputs. Using a single set of variances, e.g. the INTEG fit, we looked at how variability in the movement plan depends on readout. We found that the INTEG readout generally yielded a lower variance estimate (Fig. 5b), since it made better use of all available sensory signals (although the extent of this advantage depended on the statistical properties of the sensory transformations; Supplemental Section 2.5 and Fig. S9, online). In contrast to the idea that a single coordinate frame should dominate movement planning<sup>3-10, 15, 18</sup>, this analysis illustrates that utilizing multiple representations of a movement plan yields more reliable performance across tasks.

### Model Predictions for Previously Published Datasets

We tested the model, fit to our own dataset, on an similar, previously published dataset<sup>8</sup> containing visual target trials with an expanded range of movements (i.e., more start, and gaze locations). These data exhibit the retinal eccentricity effect, and, as above, the effect magnitude is smaller when visual feedback of the hand is available (data reproduced in Fig. 6a,b). In addition, there is a component of the reach error that correlates with the relative positions of the hand and target (data reproduced Fig. 6c,d). Our model captured all of these key features in the dataset (Fig. 6e–h).

The model explains another very different empirical result, again without additional parameter fitting. Our lab has previously shown that visual information about initial hand location is weighted more heavily when reaching to visual targets (as in VIS/FB trials here) than when reaching to proprioceptive targets (as in PROP/FB trials here)<sup>2</sup>. We proposed that

this sensory reweighting was due to the cost (e.g. variability) incurred by sensory transformations<sup>2</sup>. The present model makes this cost explicit, with quantitative predictions of the angular error that should result from artificial shifts in visual feedback. In both the empirical data and the INTEG readout predictions, visual feedback shifts had a weaker effect when reaching to proprioceptive targets than when reaching to visual targets (Fig 7a). This effect was quantified in terms of overall weighting of visual versus proprioceptive feedback (Fig. 7b), which was much greater for VIS targets than for PROP targets. In the model, this re-weighting was due to the tradeoff between the retinotopic and body-centered representations of the movement plan (Supplemental Fig. S8, online), evidenced by the fact that neither the RET nor BODY readout exhibited the effect. This result provides further support for the use of multiple representations in movement planning.

### Origins of the Gaze Bias

We have shown that a bias in the internal estimate of gaze location can account for the complex pattern of gaze-dependent reach errors we observed across trial types. We now consider several possible origins of this bias and discuss additional evidence for its presence. This bias might arise due to either a “covert” saccade plan toward the target or a shift of attention to the target. We tested these hypotheses by controlling the saccade target or the locus of attention independent of the reach target, but these manipulations did not alter the reach error pattern (Supplemental Sections 3.1, 3.2 and Fig. S10, online). Alternatively, the bias could arise from a prior expectation that reach targets tend to be foveated, reflecting the fact that eye and hand movements tend to be tightly linked<sup>31</sup>. Indeed, the model bias was implemented in this manner (see *Methods*).

While the source of the gaze bias remains undetermined, we were able to corroborate its presence using an independent measure. Specifically, we found that a visually peripheral reach target biases a subject's estimate of straight ahead, and this bias is consistent with a shift in the estimated gaze direction<sup>32</sup> towards the target (Supplemental Section 3.3 and Fig. S11, online). This perceptual effect, like the errors observed in our reach experiment, is well modeled by a Bayesian prior on gaze location centered at the target.

### Discussion

This study was aimed at testing two widely held ideas in the field of sensorimotor control: that spatial patterns of errors for a given movement reflect the underlying neural representation<sup>7-13</sup>, and a single reference frame should dominate movement planning<sup>3-6, 8-10, 15, 18, 33</sup>. We have argued that neither of these ideas is correct. First, we have shown that a single, apparently retinotopic, pattern of reach errors can be explained by a model in which multiple neural representations are used, e.g. a combination of both retinotopic and body-centered reference frames. Second, we have shown that using more than one representation confers an advantage in terms of reduced planning variability.

Spatial patterns of reach errors, especially retinotopic or gaze-centered patterns, have been cited as evidence that the neural representation for reach planning is in a particular reference frame<sup>5, 8-13, 33</sup>. We observed similar error patterns, but found that their magnitude and directionality vary with the sensory signals available for movement planning. This variation



was inconsistent with a fixed bias in a single neural representation. We showed that a bias in the transformation of target information between representations can lead to gaze-dependent error patterns in both retinotopic and body-centered representations. In fact, the retinal eccentricity effect for visual targets arose in the model only from the body-centered representation, clearly illustrating that the spatial pattern of reach errors need not be a good indicator of the underlying neural representation. More generally, we argue that when sensory signals are used in a statistically optimal manner<sup>28, 34-38</sup>, the same information is contained in multiple neural representations, and so there is no required relationship between the behavioral output and any single representation of the movement plan.

We have proposed that a biased transformation can account for gaze-dependent reaching errors, focusing on azimuthal (left-right) reach errors and azimuthal gaze shifts. When gaze is shifted vertically (up-down) a similar pattern of vertical gaze-dependent errors is observed<sup>39</sup>, qualitatively consistent with our model. We did not observe gaze-dependent effects in the radial (depth) reach errors of our subjects (Supplemental Fig. S4, online), likely because gaze was not varied in depth. When this manipulation is done, a complex pattern of gaze-dependent depth errors emerges<sup>33</sup>, presumably reflecting the complex binocular, three-dimensional geometry of the eyes<sup>29</sup>. Modeling these patterns will likely require significantly more complex model representations and transformations.

While we have focused on gaze-dependent errors, a variety of other error patterns are amenable to similar analysis. For instance, several studies exploring the effect of head rotations on reach accuracy have found errors in reaching toward the direction in which the head is oriented<sup>40, 41</sup>, a pattern consistent with a bias of the perceived midline of the head toward the direction of gaze<sup>41</sup>. Additionally, the leftward bias in reaching that we observed (Fig. 2) can be modeled as a rightward bias in the proprioceptive estimate of the right hand (Supplemental Section 1.2, online). The idiosyncratic gaze-independent error patterns exhibited by individual subjects (Supplemental Fig. S2, online) can also likely be explained by subject-specific biases on sensory variables<sup>42</sup>. All of these biases could be due to prior expectations on sensory variables or their correlations<sup>36</sup>, although alternate sources of errors are plausible, e.g. impoverished neural representations<sup>43</sup>.

We have argued that the variability of sensory transformations make it advantageous to use multiple representations for a movement plan. First, we showed that using multiple reference frames, in a weighted fashion, improves movement variability across trials (Fig. 5b). Next, we showed that only multiple reference frames in the model can account for the reweighting of visual feedback of the hand due to target type (Fig. 7)<sup>2</sup>. This is because the relative variability of the two representations, and hence their contribution to the output, depends on the sensory modality of the target. Note however, that the model does not account for all of the observed reweighting (Fig. 7). This may be explained by reweighting due to a lack of belief by the subject that the visual cursor is at their finger<sup>44</sup>. Still, our results suggest that while a single representation may simplify the flow of information<sup>3-6</sup>, it does not make optimal use of that information in movement planning.

Previous studies have reported patterns of movement errors<sup>2, 12, 26</sup> or generalization of motor learning<sup>45, 46</sup> that cannot be explained succinctly in a single reference frame. Such



error patterns often depend on the availability of sensory signals<sup>2, 12</sup>, and are explained in terms of changes in the underlying reference frame<sup>44, 46</sup> or an “intermediate” representation of movement planning<sup>26, 45</sup>. Here we provide a different explanation: movements are always represented in multiple reference frames, independent of the task, and it is the statistical reliability of these representations that determines their relative weighting.

This model is consistent with the neurophysiological literature, where a variety of spatial representations are observed across the reach planning network<sup>24, 47</sup>. Reach related areas have been found to have retinotopic coding<sup>15, 48</sup>, head- or body-centered coding<sup>19, 49</sup>, and hand- and shoulder-centered coding<sup>16, 21-23</sup>, as well as “mixed” representations<sup>14, 16-18, 21-23, 49</sup>. The two separate representations of movement plan in our model could be found in subsets of these cortical areas. Alternatively, the same computation could be performed using a heterogeneous neural population that contains both retinotopic and body-centered components, i.e. a “mixed” representation. Indeed, these implementations are at two ends of a continuum, and the physiology seems to point to a middle ground, in which all the cortical areas for reach planning make use of “mixed” representations, but the parietal cortex has a more retinotopic character and the frontal cortex is more hand or body-centered<sup>24, 47</sup>. As this study has shown, however, the ultimate answer is likely to be found not by finer assays of neural reference frames, but rather by comparing activity in these areas across tasks with different sensory information<sup>4, 14, 49</sup>.

## Materials and Methods

### Experimental setup

Subjects were seated in a virtual reality setup (Supplemental Fig. S1). The right arm rested on top of a thin (6 mm) and rigid horizontal table. The left arm remained under the table. When used as a reach target, the left index finger touched the underside the table with wrist supine. Thus, the two hands never came in contact. The location of both index fingers was monitored using an infrared tracker (Optotrak 3020, Northern Digital, Waterloo ON). Subjects' view of their arms was blocked by a mirror through which they viewed a rear-projection screen (Supplemental Fig. S1a, inset). This provided the illusion that visual objects appeared to lie in the plane of the table. The rig was enclosed in black felt and the room was darkened to minimize additional visual cues. Head movements were lightly restrained with a chin rest, and eye movements were monitored with an ISCAN Inc. (Burlington, MA) infrared eye tracker.

### Task design

Nine potential reach targets were located on the table, on a 35 cm arc centered at the point directly below the midpoint of the two eyes (Supplemental Fig. S1). The targets (8 mm radius green disks, when visual) were located at  $\pm 20^\circ$ ,  $\pm 15^\circ$ ,  $\pm 6^\circ$ ,  $\pm 2^\circ$ , and  $0^\circ$  relative to midline. Two gaze fixation points (5mm radius red disks) were located on this arc at  $\pm 10^\circ$ . Visual feedback was given with an 8 mm radius disk centered on the index finger, white for the right hand, blue for the left.

All trials consisted of four steps. 1) Subjects moved their right index finger to a fixed starting location. On feedback (FB) trials, the start location was indicated with a 10 mm radius green disk, and visual feedback of right hand was illuminated. On no-feedback (NoFB) trials, neither feedback nor target were visible, and subjects were guided to the start location using the arrow field method, which provides no feedback of absolute hand position<sup>2</sup>. 2) When the right index finger came to rest within 10 mm of the start location, one of the fixation points appeared. Subjects were required maintain fixation at the fixation point for the remainder of the trial. 3) After fixation, the reach target was specified. For VIS and VIS+PROP trials, the target disk appeared. In VIS+PROP trials, feedback of the left hand appeared, and subjects moved the left index finger to the target, and the target disk was then extinguished leaving the blue feedback disk. For PROP trials an arrow field was used to guide the unseen left hand to the unseen target location. 4) After target specification, there was a 500ms delay before an audible “go” tone was played and subjects reached to the target. On FB trials, both feedback and start disk were extinguished at the “go” tone and remained off for the rest of the trial. Finally, subjects were required to hold the final reach position for 500 ms. Subjects practiced the various trial types before beginning the experiment.

Eight subjects (two female, six male) participated in the experiment. Subjects were right-handed, had no known neural or motor deficits, had normal or corrected-to-normal vision, and gave written informed consent before participation. The experiment was divided into two sessions, performed on different days to minimize fatigue. One session contained FB trials, one NoFB trials, and session order was randomized across subjects. Each session contained six repetitions of each of the 54 conditions (3 target types  $\times$  9 targets  $\times$  2 fixation points), for a total of 324 trials (not including error trials, which were repeated). These trials were followed by trials in which gaze location was unconstrained. This set consisted of six repetitions of 9 trial conditions (3 target types  $\times$  3 targets), bringing the *total* number of trials for each session to 378. The order of presentation across conditions was randomized within each repetition.

## Data analysis

For each trial, reach endpoint was defined as the position where movement speed first fell to 5 mm/sec. Reach targets and endpoints were converted into polar coordinates about an origin located directly below the midpoint of the two eyes. *Angular reach error* is the angular difference between the endpoint and target, with positive values indicating reach endpoints to the right of the target and negative values indicating reach endpoints to the left of the target. For the plots, permutation tests, and model fitting, the angular reach errors were corrected by subtracting off the linearly interpolated free-gaze errors (separately for each subject and trial type) to minimize the effects of idiosyncratic gaze-independent errors while preserving the relationship between error, gaze, and target (see Supplemental Figs. S3, S4, and S5). The significance of gaze-dependent effects was tested by a paired permutation test of main effect of gaze location<sup>50</sup>.

## Model

Our model of reach planning describes how statistical representations of sensory inputs are used to compute a movement vector plan. Five sensory signals are potentially available, modeled as independent Gaussian probability distributions centered on the true locations,  $X$ , i.e. Gaussian likelihoods,  $N(X, \sigma^2)$ , with an isotropic covariance matrix  $\sigma^2 I$ :

$$\begin{array}{ll} \text{vision of right fingertip} & p(f_v|F_R) \sim N(F_R, \sigma_v^2) \\ \text{proprioception of right fingertip} & p(f_p|F_B) \sim N(F_B, \sigma_p^2) \\ \text{vision of target} & p(t_v|T_R) \sim N(T_R, \sigma_v^2) \\ \text{proprioception of target} & p(t_p|T_B) \sim N(T_B, \sigma_p^2) \\ \text{felt gaze position} & p(g|G) \sim N(G, \sigma_g^2) \end{array} \quad (1)$$

Here lower-case variables are sensory signals with subscripts denoting sensory modality:  $v$ , visual,  $p$ , proprioceptive. Upper-case variables are true locations with subscripts denoting reference frame:  $R$  for the retinotopic location,  $B$  for body-centered location. When a signal is unavailable in a given trial type (e.g.,  $f_v$  in NoFB trials), the likelihood is the uniform distribution. The likelihood represents *variability* in a sensory signal  $x$  given the true location  $X$ . The computations in the model, however, depend on *uncertainty* in  $X$  given  $x$ , i.e. on the posterior distributions  $p(X|x)$ . Bayes' rule relates these two distributions:

$$p(X|x) \propto p(x|X) p(X), \quad (2)$$

where  $p(X)$  represents prior information about the location. We usually assume that the prior is flat, so the posterior is proportional to the likelihood (although a non-trivial prior is described below).

All of the computations in the model make locally optimal use of the signals required for a computation, assuming that those signals are independent. Indeed, only two operations are performed by the network: signal integration and addition (or subtraction). *Integration* is the process of combining information about variable  $X$  from two signals  $x_1$  and  $x_2$ . By Bayes' rule (Equation 2) and the definition of independence, the integrated posterior is the product of the two input distributions:

$$p(X|x_1, x_2) \propto p(X|x_1) p(X|x_2). \quad (3)$$

The resulting posterior is also Gaussian, with mean and variance:

$$\mu_{X|x_1, x_2} = \left( \frac{\mu_{X|x_1}}{\sigma_{X|x_1}^2} + \frac{\mu_{X|x_2}}{\sigma_{X|x_2}^2} \right) \sigma_{X|x_1, x_2}^2, \quad \sigma_{X|x_1, x_2}^2 = \left( \frac{1}{\sigma_{X|x_1}^2} + \frac{1}{\sigma_{X|x_2}^2} \right)^{-1}. \quad (4)$$

In Equation 4 and below,  $\mu_{X|x}$  and  $\sigma_{X|x}^2$  represent the mean and variance of the distribution  $p(X|x)$ , respectively. If either input distribution is uniform, i.e. if a sensory signal is absent, the integrated posterior is equal to the other input. Note that the integrated variance is smaller than either of the input variances. The second operation, *addition*, is where the

network computes the posterior of a variable  $Z=X+Y$  from input signals  $x$  and  $y$ . In this case, the output posterior is given by convolving the inputs:

$$p(Z|y, x) = p(X|x) \otimes p(Y|y). \quad (5)$$

The result is also a Gaussian:

$$\mu_{Z|x,y} = \mu_{X|x} + \mu_{Y|y}, \quad \sigma_{Z|x,y}^2 = \sigma_{X|x}^2 + \sigma_{Y|y}^2. \quad (6)$$

Note that for both integration and addition, the output mean is a weighted sum of the input means, with either constant (unity) weights or weights that depend on the input variances.

Given the sensory signals described in Equation 1, the model first builds internal representations of fingertip and target locations, in both retinotopic and body-centered reference frames. These representations integrate all available sensory signals, requiring the transformation of “non-native” signals. In computing the retinotopic representation of target, for example, the proprioceptive signal  $t_p$  must be transformed. Since head position is fixed,  $T_R = T_B - G$ , and the transformation follows Equation 5:

$$p(T_R|t_p, g) = p(T_B|t_p) \otimes p(-G|g). \quad (7)$$

Parallel transformations convert  $f_p$  into a retinotopic representation and  $t_v$  and  $f_v$  into body-centered representations. This can yield two independent estimates of the same variable, which are then integrated according to Equation 3. The retinotopic representation of target, for example, has the posterior distribution:

$$p(T_R|t_v, t_p, g) \propto p(T_R|t_v) p(T_R|t_p, g). \quad (8)$$

Of course, in VIS or PROP trials, one of the input distributions is uniform.

The model next computes retinotopic and body-centered representations of the desired movement vector. Since the true value of the instructed movement vector is  $MV = T - F$ , these computations also follow Equation 5:

$$p(MV_U|t_v, t_p, f_v, f_p, g) = p(T_U|t_v, t_p, g) \otimes p(F_U|f_v, f_p, g), U=R, B \quad (9)$$

Note that both of the inputs to Equation 9 depend on gaze. If the same estimate of gaze is used in all transformations, then the inputs are not fully independent, and Equation 6 is only an approximate solution (see Supplemental Sections 2.2, 2.5 and Fig. S9, online).

Lastly, the model selects a planned movement vector using one of three readout schemes. The RET and BODY readouts are just the mean values of the  $MV_R$  and  $MV_B$  posteriors, respectively, from Equation 9. For the INTEG readout, the model integrates these two posteriors according to Equation 3, as if they were independent:

$$p(MV_{INTEG}|t_v, t_p, f_v, f_p, g) \propto p(MV_R|t_v, t_p, f_v, f_p, g) p(MV_B|t_v, t_p, f_v, f_p, g) \quad (10)$$

Note that these three readouts are maximum *a posteriori* (MAP) estimates given the respective posteriors. This computation is only correct when the input signals are independent (Equations 8,9), which is not always true (see Supplementary Section 2.2, online)

The final component of the model is a bias in the transformation of target position between reference frames. We model this bias as a systematic misestimation of gaze location due to a Bayesian prior. The prior takes the form of a Gaussian distribution,  $p(G) \sim N(t_{prior}, \sigma_{prior}^2)$ . The mean of the prior distribution  $t_{prior}$  is a Gaussian random variable with mean  $T_B$  and variance proportional to that of the transformed target variable. The variance of the prior,  $\sigma_{prior}^2$ , is a product of the variance of the target distribution being transformed and a scaling factor that depends on the modality of the transformed signal (see Supplemental Table S1, online). This prior only effects the gaze estimate used for transforming target locations; not the transformation of finger locations. A discussion of the statistical properties of the transformation, and their effects on model predictions, is in the Supplementary Note (Section 2.5, online).

Since the model is composed entirely of integration (Equation 3) and addition (Equation 5) operations, all expected values, including the movement vector readouts, can be written as weighted sums of the means of the initial sensory inputs, with coefficients depending only on the variances of those signals (Equations 4,6). The trial-by-trial variances of the readouts, shown in Figure 6, can be computed from these coefficients and the input variances in Equation 1.

### Model fitting

The only model parameters are the sensory variances in Equation 1 and the variance of the gaze prior,  $\sigma_{prior}^2$ . The proprioceptive variances were set *a priori* based on previously published estimates<sup>30</sup>. The variance of visual signals,  $\sigma_v^2$ , is assumed to scale linearly with the distance of the stimulus from center of gaze, and this scale factor is the first free parameter. The variance of the gaze signal,  $\sigma_g^2$  is the second. The variance of the gaze prior,  $\sigma_{prior}^2$ , is assumed to scale with variance of the target variable being transformed. Two scale factors, one for each target modality, make up the remaining free parameters. We fit these four free parameters to the average angular movement errors, after mean correction, shown in Figure 4. The model generates Cartesian movement vectors, which are then converted into polar coordinates to obtain angular errors. The fitting procedure minimized the sum square prediction error across trial conditions using the Matlab optimization toolbox (function *fmincon*; Mathworks, MA). Optimization was repeated 100 times with random initial parameter values, and the final parameters were largely insensitive to initial values.

### Modeling other data sets

The model predictions for previously published dataset (Figs. 6,7) used the model parameters fit to our own dataset with the INTEG readout. When modeling the reach errors for Buerze et. al. (Fig. 7), we translated their target array into our workspace and rescaled

the spacing of the targets to maintain the same azimuthal separation. The data for Figure 7 come from the gaze-fixed trials in Sober and Sabes<sup>2</sup>, their Supplemental Figure S3.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We would like to thank W. Pieter Medendorp and Samuel J. Sober for generously providing their data for model comparison.

**Grants:** This work was supported by the National Eye Institute (R01 EY-015679), the National Institute of Mental Health (P50 MH77970), and the McKnight Endowment Fund for Neuroscience. L.M.M.M. was supported by a graduate fellowship from the National Science Foundation.

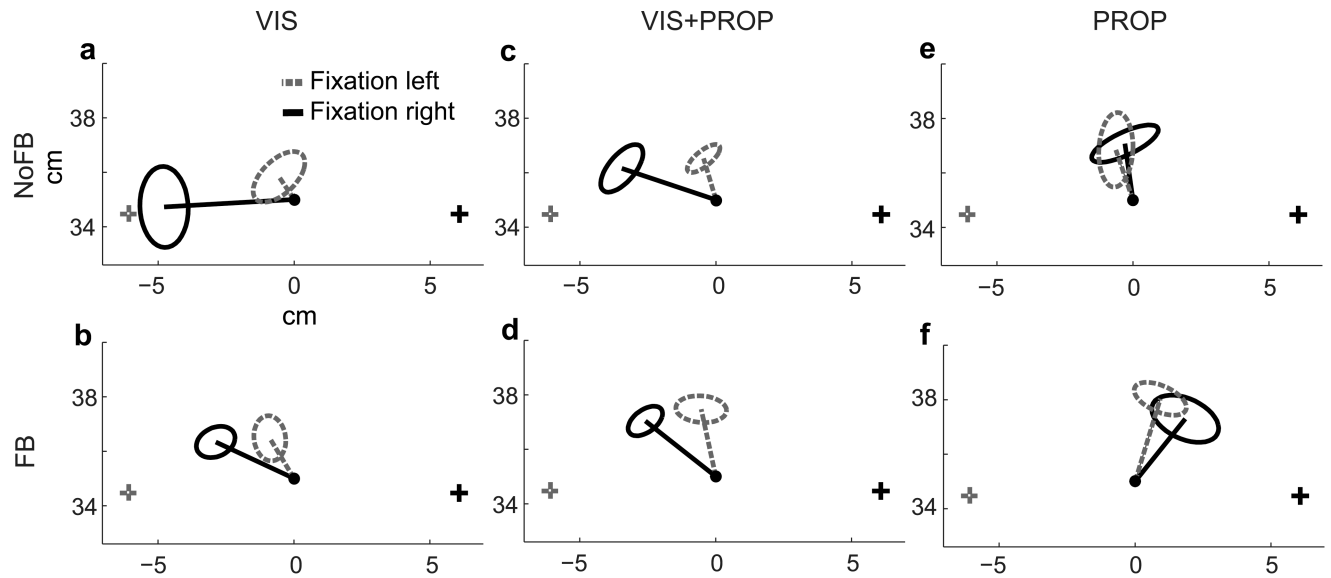
## Cited Literature

- Schlicht EJ, Schrater PR. Impact of coordinate transformation uncertainty on human sensorimotor control. *J Neurophysiol.* 2007; 97:4203–4214. [PubMed: 17409174]
- Sober SJ, Sabes PN. Flexible strategies for sensory integration during motor planning. *Nat Neurosci.* 2005; 8:490–497. [PubMed: 15793578]
- Buneo CA, Andersen RA. The posterior parietal cortex: sensorimotor interface for the planning and online control of visually guided movements. *Neuropsychologia.* 2006; 44:2594–2606. [PubMed: 16300804]
- Cohen YE, Andersen RA. A common reference frame for movement plans in the posterior parietal cortex. *Nat Rev Neurosci.* 2002; 3:553–562. [PubMed: 12094211]
- Engel KC, Flanders M, Soechting JF. Oculocentric frames of reference for limb movement. *Archives italiennes de biologie.* 2002; 140:211–219. [PubMed: 12173524]
- Lacquaniti F, Caminiti R. Visuo-motor transformations for arm reaching. *Eur J Neurosci.* 1998; 10:195–203. [PubMed: 9753127]
- Soechting JF, Flanders M. Sensorimotor representations for pointing to targets in three-dimensional space. *J Neurophysiol.* 1989; 62:582–594. [PubMed: 2769349]
- Beurze SM, Van Pelt S, Medendorp WP. Behavioral reference frames for planning human reaching movements. *J Neurophysiol.* 2006; 96:352–362. [PubMed: 16571731]
- Henriques DY, Klier EM, Smith MA, Lowy D, Crawford JD. Gaze-centered remapping of remembered visual space in an open-loop pointing task. *J Neurosci.* 1998; 18:1583–1594. [PubMed: 9454863]
- Pouget A, Ducom JC, Torri J, Bavelier D. Multisensory spatial representations in eye-centered coordinates for reaching. *Cognition.* 2002; 83:B1–11. [PubMed: 11814488]
- McIntyre J, Stratta F, Lacquaniti F. Viewer-centered frame of reference for pointing to memorized targets in three-dimensional space. *J Neurophysiol.* 1997; 78:1601–1618. [PubMed: 9310446]
- Carrozzo M, McIntyre J, Zago M, Lacquaniti F. Viewer-centered and body-centered frames of reference in direct visuomotor transformations. *Exp Brain Res.* 1999; 129:201–210. [PubMed: 10591894]
- McIntyre J, Stratta F, Lacquaniti F. Short-term memory for reaching to visual targets: psychophysical evidence for body-centered reference frames. *J Neurosci.* 1998; 18:8423–8435. [PubMed: 9763485]
- Avillac M, Deneve S, Olivier E, Pouget A, Duhamel JR. Reference frames for representing visual and tactile locations in parietal cortex. *Nat Neurosci.* 2005; 8:941–949. [PubMed: 15951810]
- Batista AP, Buneo CA, Snyder LH, Andersen RA. Reach plans in eye-centered coordinates. *Science.* 1999; 285:257–260. [PubMed: 10398603]
- Batista AP, et al. Reference frames for reach planning in macaque dorsal premotor cortex. *J Neurophysiol.* 2007; 98:966–983. [PubMed: 17581846]

17. Battaglia-Mayer A, et al. Eye-hand coordination during reaching. II. An analysis of the relationships between visuomanual signals in parietal cortex and parieto-frontal association projections. *Cereb Cortex*. 2001; 11:528–544. [PubMed: 11375914]
18. Buneo CA, Jarvis MR, Batista AP, Andersen RA. Direct visuomotor transformations for reaching. *Nature*. 2002; 416:632–636. [PubMed: 11948351]
19. Lacquaniti F, Guigon E, Bianchi L, Ferraina S, Caminiti R. Representing spatial information for limb movement: role of area 5 in the monkey. *Cereb Cortex*. 1995; 5:391–409. [PubMed: 8547787]
20. Medendorp WP, Goltz HC, Vilis T, Crawford JD. Gaze-centered updating of visual space in human parietal cortex. *J Neurosci*. 2003; 23:6209–6214. [PubMed: 12867504]
21. Pesaran B, Nelson MJ, Andersen RA. Dorsal premotor neurons encode the relative position of the hand, eye, and goal during reach planning. *Neuron*. 2006; 51:125–134. [PubMed: 16815337]
22. Wu W, Hatsopoulos N. Evidence against a single coordinate system representation in the motor cortex. *Exp Brain Res*. 2006; 175:197–210. [PubMed: 16775704]
23. Wu W, Hatsopoulos NG. Coordinate system representations of movement direction in the premotor cortex. *Exp Brain Res*. 2007; 176:652–657. [PubMed: 17180398]
24. Burnod Y, et al. Parieto-frontal coding of reaching: an integrated framework. *Exp Brain Res*. 1999; 129:325–346. [PubMed: 10591906]
25. Caminiti R, Ferraina S, Mayer AB. Visuomotor transformations: early cortical mechanisms of reaching. *Curr Opin Neurobiol*. 1998; 8:753–761. [PubMed: 9914239]
26. Carrozzo M, Lacquaniti F. A hybrid frame of reference for visuo-manual coordination. *Neuroreport*. 1994; 5:453–456. [PubMed: 8003674]
27. Bock O. Contribution of retinal versus extraretinal signals towards visual localization in goal-directed movements. *Exp Brain Res*. 1986; 64:476–482. [PubMed: 3803485]
28. Knill DC, Pouget A. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci*. 2004; 27:712–719. [PubMed: 15541511]
29. Blohm G, Crawford JD. Computations for geometrically accurate visually guided reaching in 3-D space. *J Vis*. 2007; 7:1–22. 4.
30. van Beers RJ, Sittig AC, Denier van der Gon JJ. The precision of proprioceptive position sense. *Exp Brain Res*. 1998; 122:367–377. [PubMed: 9827856]
31. Land MF, Hayhoe M. In what ways do eye movements contribute to everyday activities? *Vision Res*. 2001; 41:3559–3565. [PubMed: 11718795]
32. Balslev D, Miall RC. Eye position representation in human anterior parietal cortex. *J Neurosci*. 2008; 28:8968–8972. [PubMed: 18768690]
33. Van Pelt S, Medendorp WP. Updating target distance across eye movements in depth. *J Neurophysiol*. 2008; 99:2281–2290. [PubMed: 18353912]
34. Ernst MO, Banks MS. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*. 2002; 415:429–433. [PubMed: 11807554]
35. Helbig HB, Ernst MO. Optimal integration of shape information from vision and touch. *Exp Brain Res*. 2007; 179:595–606. [PubMed: 17225091]
36. Kording KP, Wolpert DM. Probabilistic mechanisms in sensorimotor control. *Novartis Found Symp*. 2006; 270:191–198. discussion 198–202, 232–197. [PubMed: 16649715]
37. van Beers RJ, Sittig AC, Gon JJ. Integration of proprioceptive and visual position-information: An experimentally supported model. *J Neurophysiol*. 1999; 81:1355–1364. [PubMed: 10085361]
38. Yuille, AL.; Bulthoff, HH. Perception as Bayesian Inference. Cambridge University Press; 1996. Bayesian decision theory and psychophysics; p. 123–136. Published by
39. Henriques DY, Crawford JD. Direction-dependent distortions of retinocentric space in the visuomotor transformation for pointing. *Exp Brain Res*. 2000; 132:179–194. [PubMed: 10853943]
40. Henriques DY, Crawford JD. Role of eye, head, and shoulder geometry in the planning of accurate arm movements. *J Neurophysiol*. 2002; 87:1677–1685. [PubMed: 11929889]
41. Lewald J, Ehrenstein WH. Visual and proprioceptive shifts in perceived egocentric direction induced by eye-position. *Vision Res*. 2000; 40:539–547. [PubMed: 10820612]

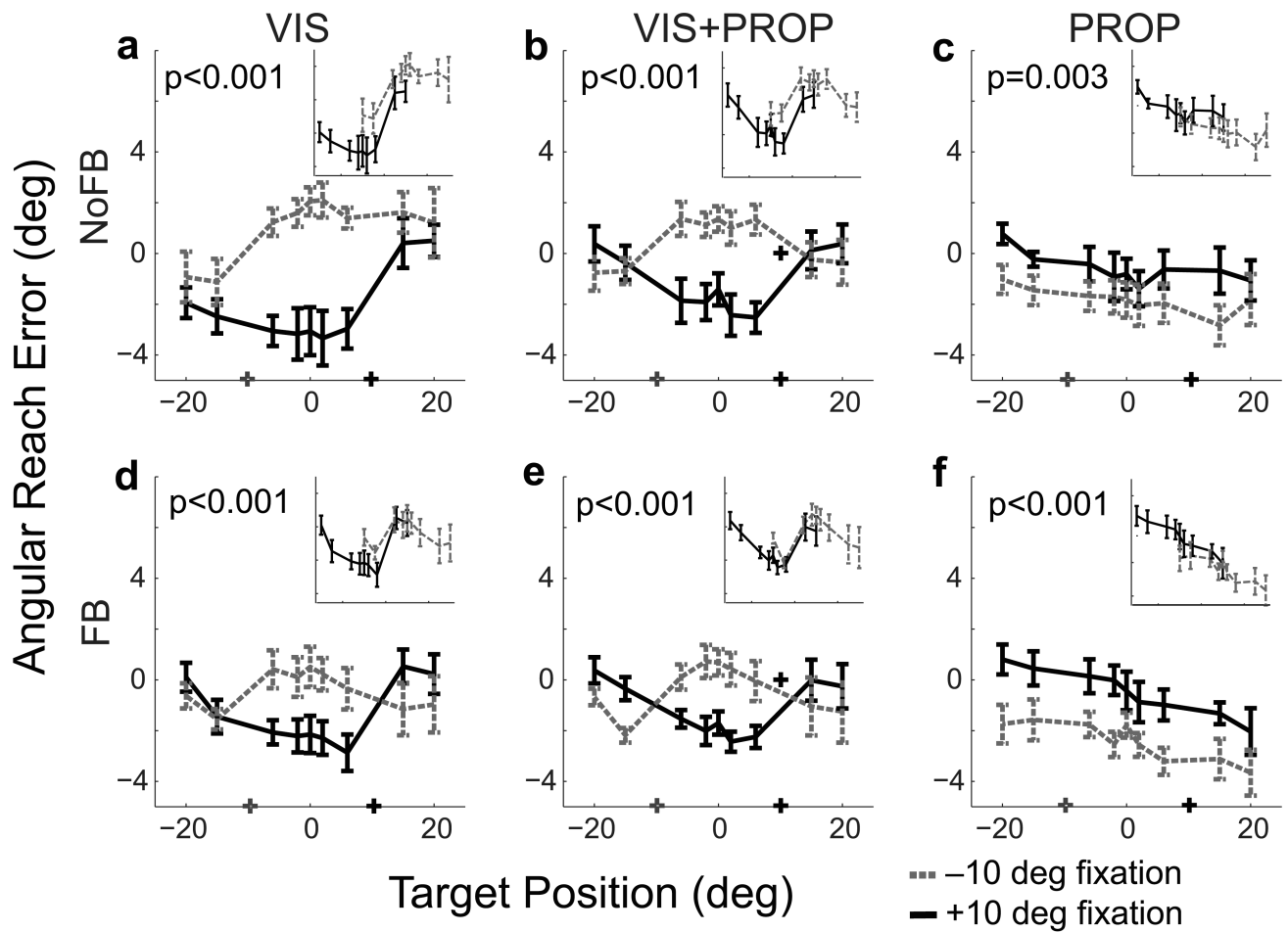


42. Vindras P, Desmurget M, Prablanc C, Viviani P. Pointing errors reflect biases in the perception of the initial hand position. *J Neurophysiol.* 1998; 79:3290–3294. [PubMed: 9636129]
43. Soechting JF, Flanders M. Errors in pointing are due to approximations in sensorimotor transformations. *J Neurophysiol.* 1989; 62:595–608. [PubMed: 2769350]
44. Kording, KP.; Tenenbaum, JB. NIPS. MIT Press; 2006. Causal inference in multisensory integration; p. 737-744.
45. Ahmed AA, Wolpert DM, Flanagan JR. Flexible representations of dynamics are used in object manipulation. *Curr Biol.* 2008; 18:763–768. [PubMed: 18485709]
46. Kluzik J, Diedrichsen J, Shadmehr R, Bastian AJ. Reach adaptation: what determines whether we learn an internal model of the tool or adapt the model of our arm? *J Neurophysiol.* 2008; 100:1455–1464. [PubMed: 18596187]
47. Battaglia-Mayer A, Caminiti R, Lacquaniti F, Zago M. Multiple levels of representation of reaching in the parieto-frontal network. *Cereb Cortex.* 2003; 13:1009–1022. [PubMed: 12967918]
48. Cohen YE, Andersen RA. Reaches to sounds encoded in an eye-centered reference frame. *Neuron.* 2000; 27:647–652. [PubMed: 11055445]
49. Mullette-Gillman OA, Cohen YE, Groh JM. Eye-centered, head-centered, and complex coding of visual and auditory targets in the intraparietal sulcus. *J Neurophysiol.* 2005; 94:2331–2352. [PubMed: 15843485]
50. Good, PI. *Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses.* 2nd. New York: Springer-Verlag; 2000.



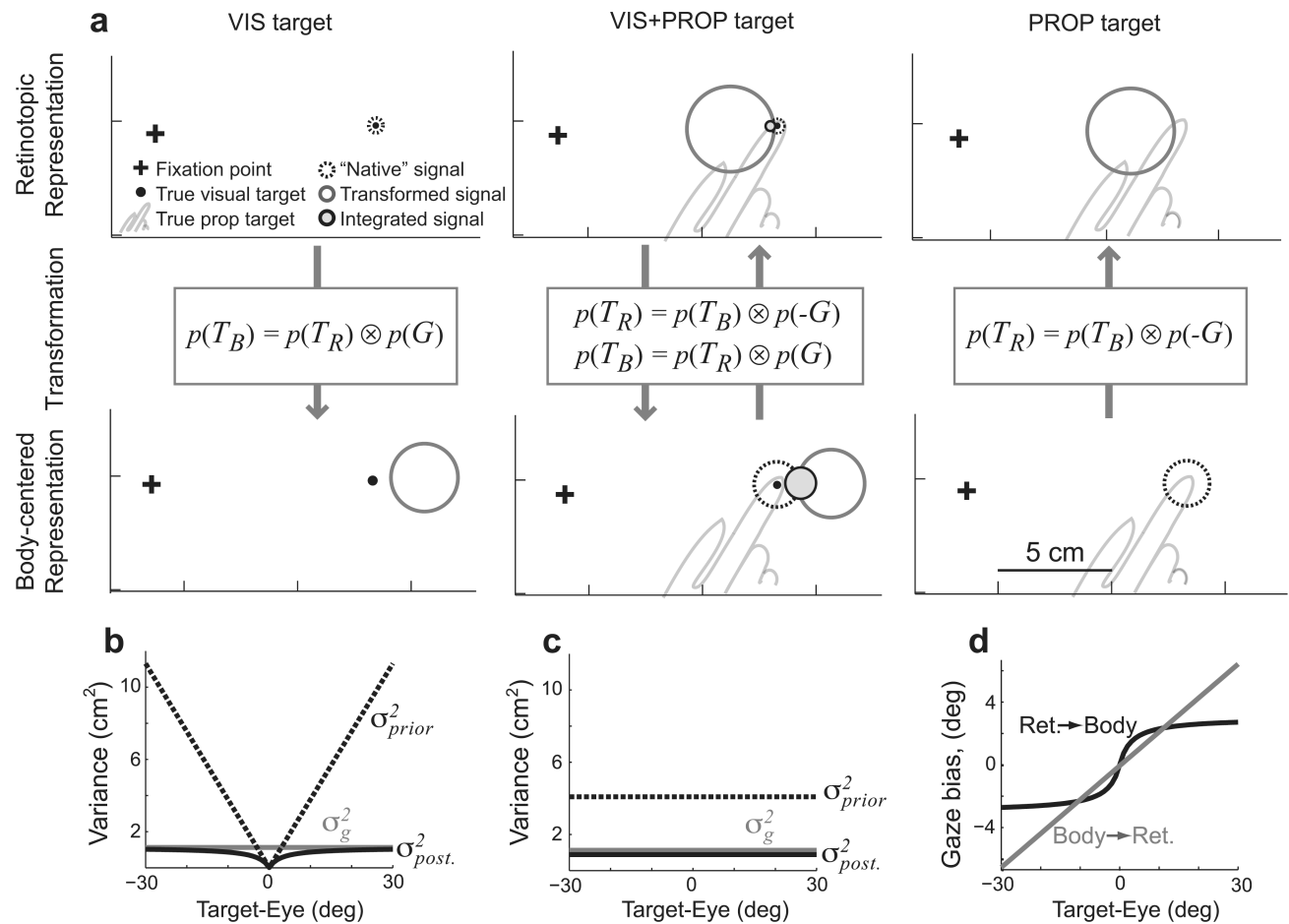
**Figure 1.**

Reach errors at the center target for an example subject for all trial types. Lines indicate mean reach error for each gaze-position, ellipses represent standard deviation, + fixation points, • reach targets. The origin (not shown) is located directly below midpoint of the eyes.

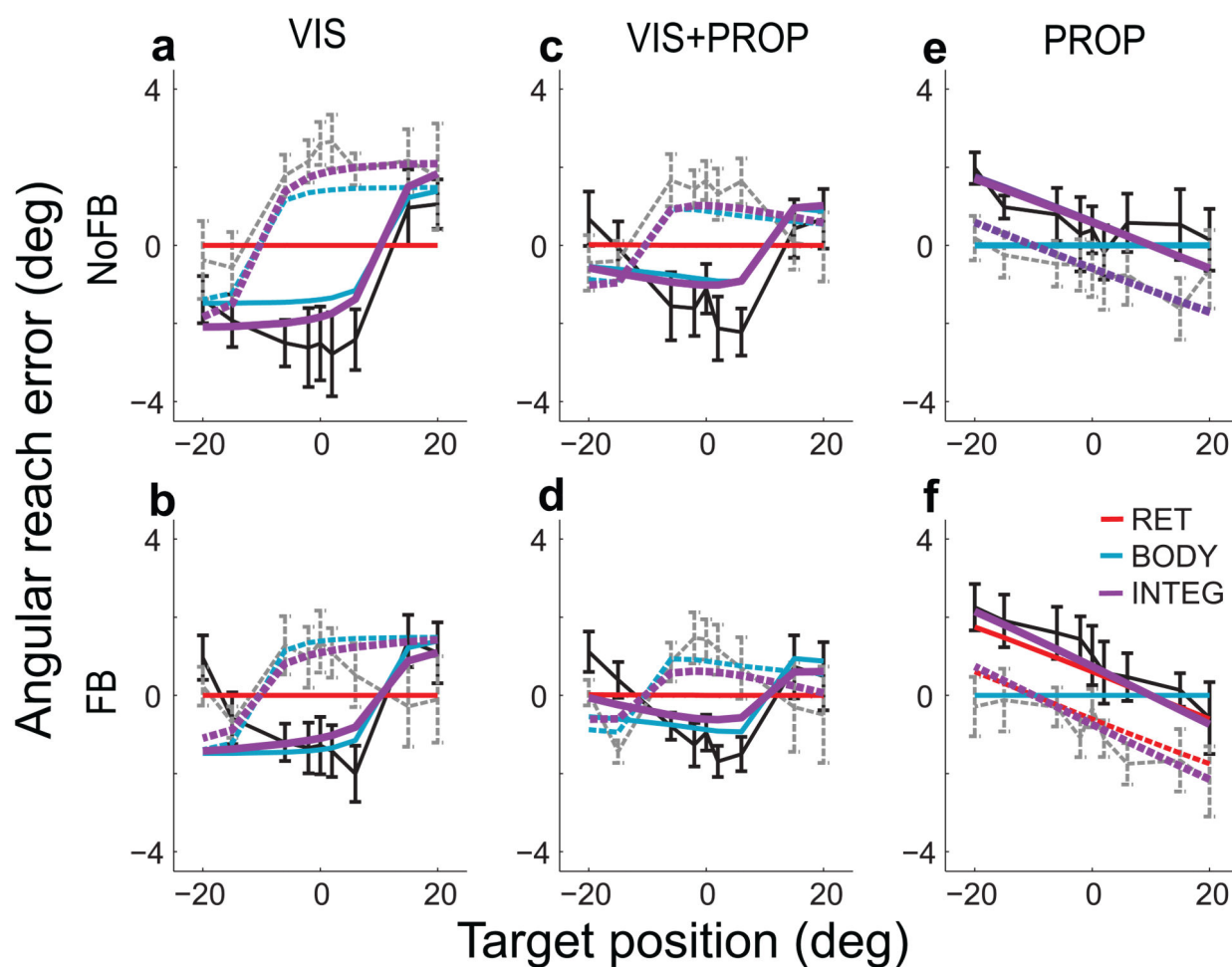


**Figure 2.**

Average angular reach error across subjects for each trial condition. Negative values indicate reach endpoints to the left of target, positive values indicate reach endpoints to the right of target. Before averaging, linearly interpolated gaze-free errors for each trial type were subtracted (Supplemental Section 1.2, and Fig. S4, online). Error bars indicate standard errors. p-values determined with a paired permutation test<sup>50</sup>.

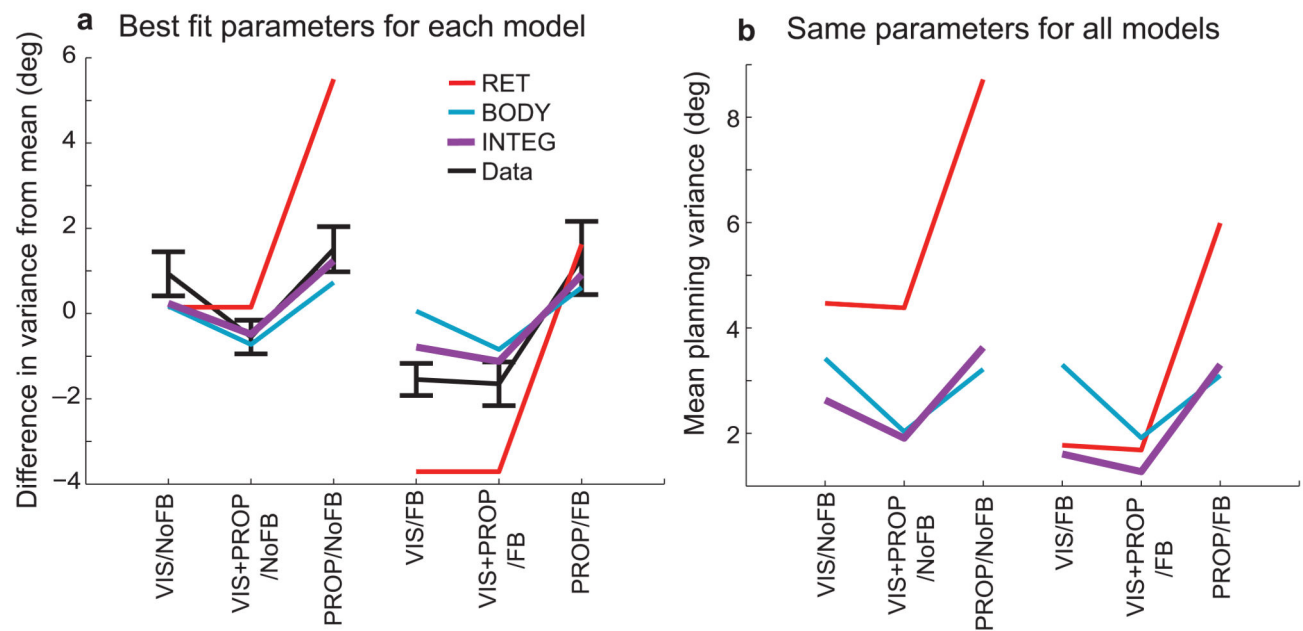
**Figure 3.**

The bias and variance injected by sensory transformations in the model. a) Examples of the bias and variance that arise during transformation of target information. Circles represent the posterior distribution of target location (95%-confidence limits). Transformed signals have greater uncertainty and are biased either towards (retinotopic) or away from (body-centered) the gaze location. In VIS+PROP trials, the “native” and transformed signals are integrated into a combined estimate of location (filled circles). b,c) Posterior variance in estimated gaze location (i.e., “transformation variance”, black), gaze likelihood variance (grey), and variance of gaze prior (dashed) (see *Methods*) for b) transformation from retinotopic to body-centered space, and c) transformation from body-centered to retinotopic space. d) Bias in eye position estimate used in sensory transformations. Values in all panels computed from INTEG model fit.

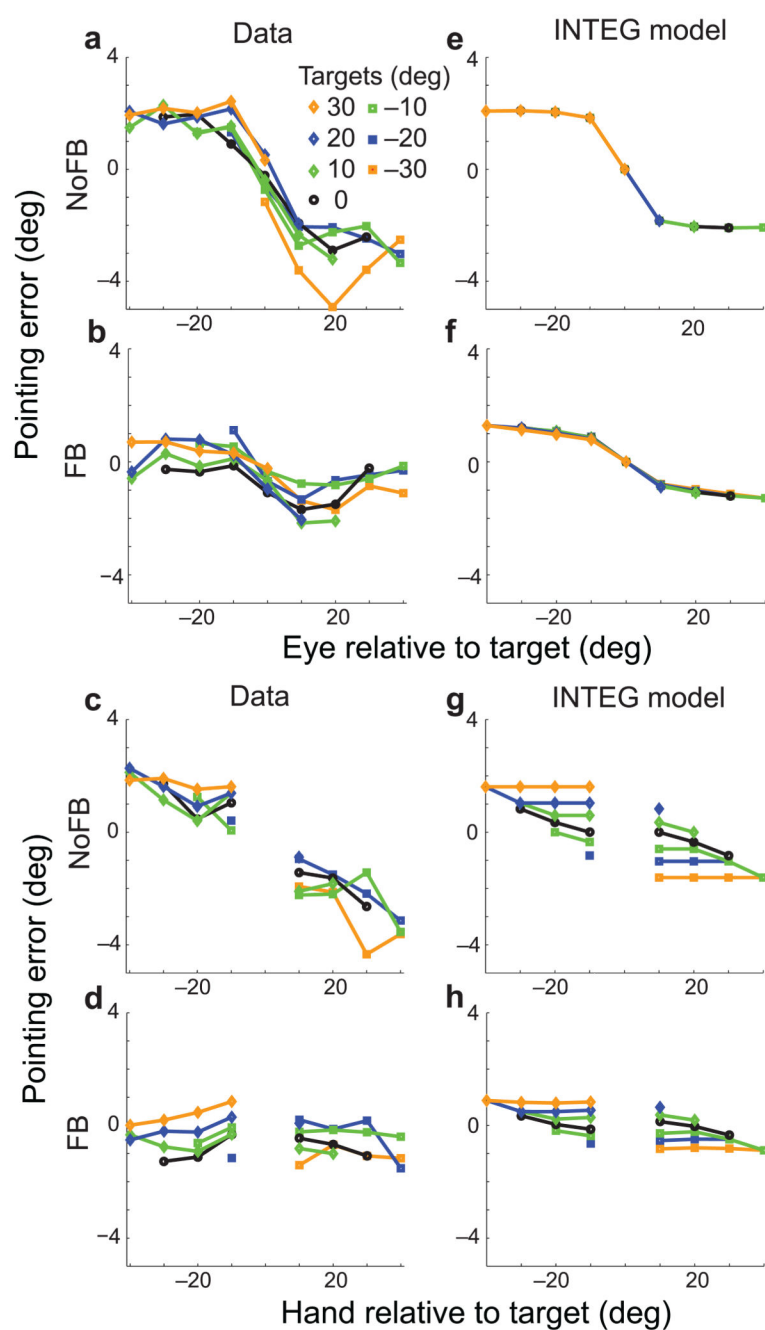


**Figure 4.**

Model fits for gaze-dependent reach errors. Black and gray lines show mean (standard error) errors across subjects for each trial condition, after subtracting the overall mean separately for each of the six trial types. Colored lines show best-fit model predictions. Solid lines, gaze right; dashed lines, gaze left.

**Figure 5.**

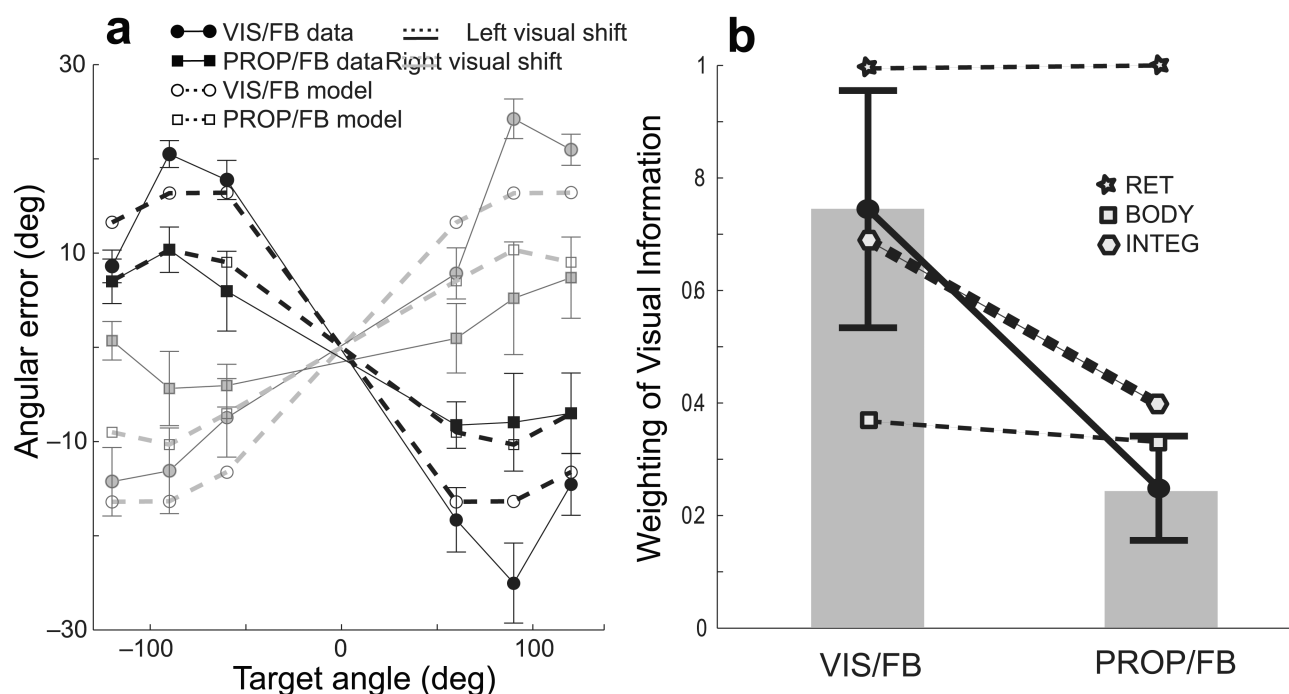
Reach variability. a) Trial-type differences in angular reach variance. Black lines represent mean (standard error) reach variance across subjects. Values are derived from the average variance across subjects and trial conditions within each trial type, with the overall mean subtracted. Colored lines represent model predictions with each readouts fit parameters. b) Mean variability of each model readout scheme, as a function of trial type, using fixed model parameters (INTEG fit).



**Figure 6.**

Model predictions for data from Beurze et. al.8 a,b,e,f) Gaze dependent pointing error. c,d,g,h) Pointing error as a function of initial hand position relative to the target. a–d) average pointing error from Beurze et al. e–h) INTEG model predictions of pointing error, with parameters fit to our data.





**Figure 7.**

Changes in sensory weighting with target modality. a) Mean angular error induced by artificial shifts in the visual feedback of the hand prior to movement onset. Data from Sober and Sabes2. Error bars represent standard error. Model predictions use the INTEG readout, parameters fit to our data. b) Relative weighting of visual vs. proprioceptive information about initial hand position in movement planning for reaches to VIS and PROP targets. Error bars represent standard deviation across subjects. Colored lines show model predictions for each readout scheme.